

Analisis Sentimen Terhadap Ulasan Restoran Menggunakan Metode *Multi Layer Perceptron*

Septian Fadillah¹, Wina Witanti², Edvin Ramadhan³

Fakultas Sains Dan Informatika/Program Studi Informatika
Universitas Jenderal Achmad Yani
Cimahi, Indonesia

e-mail: ¹sfadillah46@gmail.com, ²witanti@gmail.com, ³edvinr@gmail.com

Correspondence : e-mail: sfadillah46@gmail.com

Diajukan: 15 Agustus 2024; Direvisi: 23 Agustus 2024; Diterima: 25 Agustus 2024

Abstrak

Reputasi restoran dipengaruhi ulasan yang diberikan oleh pengunjung kepada restoran terkait. Faktor yang mempengaruhi ulasan restoran salah satunya rasa makanan, pelayanan, kebersihan, harga, dan lainnya. Rata-rata orang akan berkunjung ke suatu restoran melalui ulasan dari pengunjung sebelumnya. Dalam konteks ulasan restoran, klasifikasi sentimen analisis dapat membantu pemilik restoran atau pengelola untuk memahami pandangan pelanggan tentang pengalaman mereka di restoran tersebut. Pada penelitian sebelumnya klasifikasi ulasan sudah pernah dilakukan dengan menggunakan metode yang berbeda menggunakan metode Naïve Bayes. pada penelitian tersebut memiliki tingkat akurasi 86,50% menjadi 90.50%. Pemanfaatan sentence embedding mampu mengubah kalimat menjadi vektor numerik menggunakan Word2Vec kemudian setelahnya dilakukan proses klasifikasi dengan menggunakan Multi Layer Perceptron (MLP) yang menghasilkan dua kelas yaitu sentiment positif, negative .

Kata kunci: ulasan; sentiment analysis; Word2Vec; Multi Layer Perceptron, Restoran

Abstract

Restaurant reputation is influenced by reviews given by visitors to related restaurants. Factors that influence restaurant reviews include food taste, service, cleanliness, price, and others. The average person will visit a restaurant through reviews from previous visitors. In the context of restaurant reviews, sentiment analysis classification can help restaurant owners or managers to understand customer views about their experience at the restaurant. In previous research, the classification of reviews has been carried out using different methods using the Naïve Bayes method. in that study it had an accuracy rate of 86.50% to 90.50%. Utilization of sentence embedding is able to convert sentences into numeric vectors using Word2Vec then after that the classification process is carried out using Multi Layer Perceptron (MLP) which produces two classes, namely positive, negative sentiment.

Keywords: reviews; sentiment analysis; Word2Vec; Multi Layer Perceptron, Restaurant.

1. Pendahuluan

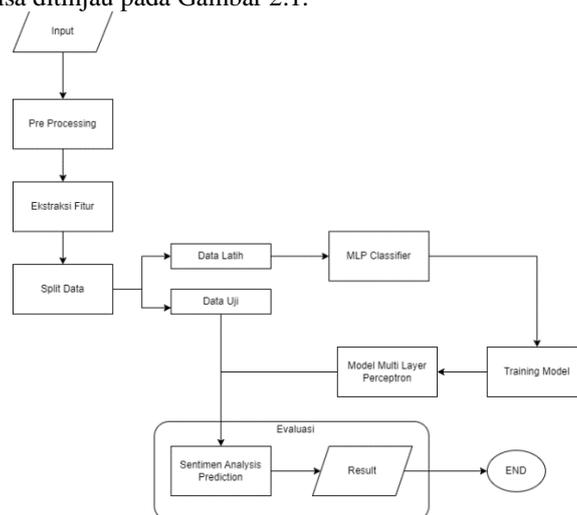
Restoran adalah tempat yang menyajikan makanan dan minuman kepada pelanggan yang datang untuk makan atau untuk berkumpul setelah bekerja. Sebuah restoran dapat menawarkan berbagai jenis makanan, termasuk makanan regional, internasional, makanan cepat saji dan hidangan khusus. Restoran menyediakan pilihan pengiriman makanan ke rumah pelanggan atau tempat kerja.[1] Reputasi restoran dipengaruhi ulasan yang diberikan oleh pengunjung kepada restoran terkait. Faktor yang mempengaruhi ulasan restoran salah satunya rasa makanan, pelayanan, kebersihan, harga, dan lainnya. Rata-rata orang akan berkunjung ke suatu restoran melalui ulasan dari pengunjung sebelumnya. Penelitian sebelumnya telah melakukan penelitian mengenai ulasan restoran *online*[2]. Penelitian lainnya melakukan ulasan produk yang disukai dan tidak disukai untuk restoran Indonesia[3]. Namun, dalam menentukan suatu pilihan memiliki pertimbangan. Oleh karenanya untuk membantu meyakinkan hasil ulasan, maka akhir-akhir ini sentiment analysis menjadi salah satu proses komputasi yang dapat membantu menentukan suatu hal. Penelitian sebelumnya *sentiment analysis* digunakan untuk menganalisis ulasan pelanggan di berbagai restoran seluruh karachi(salah satu kota di Pakistan)[4].

Tahapan mengklasifikasikan ulasan restoran dapat memanfaatkan teknik *sentence embedding*, yang merupakan suatu teknik dalam *natural language processing* (NLP) untuk mengubah kalimat menjadi vektor numerik. Penelitian sebelumnya telah melakukan evaluasi terhadap proses *sentence embedding* menggunakan metode *Bag-of-Words*, Dengan *Word2Vec* dengan merubah kalimat menjadi vektor[5]. Penggunaan metode machine learning untuk klasifikasi ulasan restoran salah satunya yaitu *Multi Layer Perceptron*(MLP) dengan memanfaatkan data yang berkaitan dengan ulasan dari pelanggan. MLP adalah sebuah jaringan saraf tiruan yang terdiri dari beberapa lapisan (*layer*) yang saling terhubung terdiri dari *input layer*, *hidden layer*, dan *output layer*. Penelitian sebelumnya dengan menggunakan Multilayer Perceptron untuk mengklasifikasikan ulasan terhadap restoran Thailand memiliki akurasi 93,5% [6] Penelitian Sebelumnya Saat dikombinasikan dengan *FastText* dalam penelitian tersebut memberikan kontribusi dalam meningkatkan kinerja model SVM yang menjadi metode klasifikasi dalam penelitian tersebut dengan mencapai akurasi 82,65%[7] Penelitian sebelumnya dengan menggunakan kombinasi *multilayer perceptron* dan *word2vect* dalam penanganan kebijakan penggunaan sertifikat vaksin memiliki akurasi 60% [8].

Berdasarkan dari uraian di atas, maka penelitian ini melakukan klasifikasi ulasan restoran dengan memanfaatkan *Word2Vec* yang digunakan dalam pengubahan kalimat menjadi vektor dan *Multi Layer Perceptron* sebagai metode klasifikasi dimana akan memperhatikan pengaruh konfigurasi parameter dalam *Multi Layer Perceptron* dalam pengaruh akurasi yang didapat.

2. Metode Penelitian *Multi Layer Peceptron*

Data Mining adalah proses menemukan korelasi baru bermakna, pola dan tren dengan memilah-milah sejumlah besar data yang disimpan dalam repositori, akan dilakukan teknologi pengenalan pola serta statistic. Penelitian ini akan dilakukan melalui beberapa tahapan penelitian menggunakan metode *Multi Layer Perceptron* yang bisa ditinjau pada Gambar 2.1.



Gambar 1 Metodologi Penelitian

Proses Penelitian terdiri dari 7 tahapan yaitu

1. Pengumpulan *Dataset*

Data yang digunakan merupakan data publik berjumlah 1.018 data dari situs Kaggle (<https://www.kaggle.com/code/apekshakom/sentiment-analysis-of-restaurant-reviews/input>)

2. Ekstraksi Fitur

Word2Vect adalah sebuah teknik dalam pemrosesan bahasa *alami* yang digunakan untuk menghasilkan representasi vektor dari kata-kata. Metode ini, embedding kata digunakan untuk menggambarkan kata-kata secara numerik. Vektor yang mampu mengkodekan kata-kata yang memiliki makna yang sama cenderung berdekatan dalam ruang vektor, menunjukkan bahwa kata-kata tersebut memiliki makna yang serupa

3. Split Data

Sebelum klasifikasi dilakukan data yang sudah diberi label dikelompokkan ke dalam dua bagian: data pelatihan, pengujian. Data pelatihan dipergunakan untuk mengembangkan model, penggunaan data pengujian untuk mengukur dan menguji akurasi model

4. MLP Classifier

Multi Layer Perceptron (MLP) adalah jenis jaringan saraf tiruan yang terdiri dari beberapa lapisan tersembunyi yang saling terhubung. Setiap lapisan tersembunyi terdiri dari banyak neuron yang saling terhubung Model Multi Layer Perceptron.

5. Tahap Evaluasi

Evaluasi ini membantu dalam mengukur performa model secara objektif dan memberikan dasar untuk iterasi dan perbaikan atas data uji untuk memastikan akurasi dan efektivitas model dalam memahami dan mengklasifikasikan sentimen dari teks.

6. Sentimen Analysis Prediction

untuk memastikan akurasi dan efektivitas model dalam memahami dan mengklasifikasikan sentimen dari teks.

2.1. Multi Layer Perceptron

Multi Layer Perceptron (MLP) adalah jenis jaringan saraf tiruan yang terdiri dari beberapa lapisan tersembunyi yang saling terhubung[9]. Setiap lapisan tersembunyi terdiri dari banyak neuron yang saling terhubung seperti pada gambar 2.2. MLP digunakan dalam pembelajaran mesin untuk menyelesaikan tugas-tugas yang kompleks dengan melakukan klasifikasi dan regresi pada data input dengan menggunakan teknik-teknik seperti aktivasi Logistic(2), ReLu(3), Tanh(1) yang memiliki perhitungan untuk menentukan nilai kedalam output layer.[10]

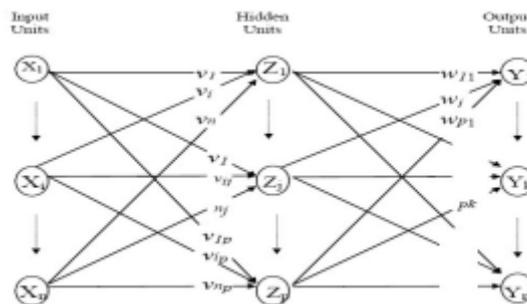
$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{1}$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

$$ReLU(x) = \max(0, x) \tag{3}$$

Secara prinsip MLP dapat disimpulkan sebagai berikut :

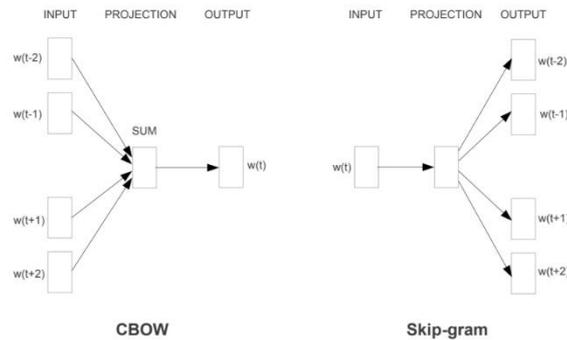
- 1 Setiap neuron yang dimiliki jaringan mempunyai fungsi aktivasi non linier dari kelas.
- 2 MLP terdiri dari satu atau lebih hidden layer .
- 3 Sinyal menembus melalui jaringan, dari input ke output , dengan arah lurus, layer ke layer
- 4 Jaringan memiliki connectivity yang tinggi (synaptic).



Gambar 2 Multi Layer Perceptron

2.2. Word2Vect

Word2Vect adalah sebuah teknik dalam pemrosesan bahasa alami yang digunakan untuk menghasilkan representasi vektor dari kata-kata. Metode ini, embedding kata digunakan untuk menggambarkan kata-kata secara numerik. Vektor yang mampu mengkodekan kata-kata yang memiliki makna yang sama cenderung berdekatan dalam ruang vektor, menunjukkan bahwa kata-kata tersebut memiliki makna yang serupa[3].



Gambar 3 Word2Vect

Word2Vect terdiri dari dua model yang berbeda, yaitu *continuous bag of words* (CBOW) dan skip-gram. CBOW bertujuan untuk menggunakan kata-kata untuk memprediksi kata-kata yang berdekatan, sedangkan skip-gram bertujuan untuk menggunakan kata-kata untuk memprediksi kata-kata yang berdekatan dengannya. Word2Vect berperan penting dalam menghasilkan representasi yang akurat dari kata-kata dalam proses *word embeddings*, yang nantinya dapat digunakan dalam menganalisis data, memahami makna kata-kata, atau bahkan dalam pengembangan model-model *deep learning* seperti yang dilakukan dalam penelitian.[11]

3. Hasil dan Pembahasan

Hasil uji coba ini diharapkan dapat melihat pentingnya konfigurasi parameter yang dilakukan untuk mendapatkan konfigurasi parameter terbaik dalam penggunaan *Multi Layer Perceptron* sebagai metode klasifikasi untuk mendapatkan hasil akurasi yang optimal.

3.1. Pengujian Parameter Hidden Layer

Hidden layer atau lapisan tersembunyi berfungsi sebagai komponen krusial yang terletak di antara lapisan input dan lapisan output. *Hidden layer* bertugas mengekstrak dan mempelajari fitur kompleks dari data input, serta memperkenalkan non-linearitas melalui fungsi aktivasi yang memungkinkan model untuk menangkap pola yang lebih rumit. Setiap neuron di hidden layer menerima input dari lapisan sebelumnya, mengalikannya dengan bobot, menambahkan bias, dan menerapkan fungsi aktivasi untuk menghasilkan output yang diteruskan ke lapisan berikutnya. Hasil dari pengujian dapat dilihat pada tabel 1.

Tabel 1 Pengujian Hidden Layer

Konfigurasi <i>Hidden Layer</i>	Akurasi Pelatihan	Akurasi Pengujian
(1024,),(50,),(50,100,50), (48,)	78.38%	82.35%
(50,),(100,),(100,50),(50)	96.15%	97.39%
(1000,)(50,),(100,50)	87.96	90.85
(1000,),(50)	59.72%	61.44%

Perbedaan akurasi pada pengujian MLP dengan konfigurasi *hidden layer* yang berbeda mencerminkan dampak dari jumlah neuron, struktur lapisan, serta potensi *overfitting* atau *underfitting*. Untuk mencapai kinerja yang optimal, penting untuk menyeimbangkan jumlah neuron dan lapisan, serta mempertimbangkan teknik regularisasi dan optimisasi yang sesuai dengan karakteristik data dan tugas. Maka dari pengujian parameter *hidden layer* yang sudah dilakukan penggunaan parameter yang digunakan dalam penelitian ini ada pada pengujian kedua dengan parameter (50,),(100,),(100,50),(50).

1. (50,): Model ini memiliki satu lapisan tersembunyi dengan 50 neuron.
2. (100,): Model ini memiliki satu lapisan tersembunyi dengan 100 neuron.
3. (100, 50): Model ini memiliki dua lapisan tersembunyi, yang pertama dengan 100 neuron dan yang kedua dengan 50 neuron.
4. (50): Model ini memiliki satu lapisan tersembunyi dengan 50 neuron.

3.2. Pengujian Aktifasi

Fungsi aktivasi memainkan peran penting dalam menentukan bagaimana model memproses input dan menghasilkan output. Fungsi aktivasi digunakan untuk memperkenalkan non-linearitas dalam model, memungkinkan jaringan saraf untuk mempelajari hubungan yang kompleks dan non-linear antara fitur input dan output. Pengujian menggunakan model parameter *hidden layer* terbaik kemudian menggunakan pengujian dari beberapa aktivasi, Hasil pengujian dapat dilihat pada tabel 2.

Tabel 2 Tabel Konfigurasi Aktifasi

Konfigurasi Aktifasi	Akurasi Pelatihan	Akurasi Pengujian
Tanh (Tangens Hyperbolik)	58.64%	62.09%
Logistic (Simoid)	58.64%	50.98%
Relu (Rectified Linear Unit)	60.91%	62.09%
Tanh, Logistic, dan Relu	96.15%	97.39%

Dalam pengujian konfigurasi parameter *Multi Layer Perceptron* dengan konfigurasi hidden layer yang sama tetapi fungsi aktivasi yang berbeda menunjukkan pentingnya memilih fungsi aktivasi yang tepat untuk model.

3.2.1. Pengujian Parameter Learning_Rate

Learning rate adalah parameter penting yang mempengaruhi seberapa cepat atau lambat model belajar dari data dapat dilihat pada tabel 3.

Tabel 3 Tabel Konfigurasi Hasil Learning_rate

Konfigurasi Learning_rate	Akurasi Pelatihan	Akurasi Pengujian
Constant	59.72%	61.44%
Adaptive	60.32%	63.73%
Adaptive Constant	96.15%	97.39%

Penggabungan learning rate antara adaptive dan constant dimana menggabungkan dua metode dengan harapan dapat mengoptimalkan stabilitas dari learning rate constant dan penyesuaian dinamis dari *learning rate adaptive*.

4. Kesimpulan

Berdasarkan pengujian yang telah dilakukan pada analisis sentiment terhadap ulasan restoran, menggunakan algoritma MLP dengan *Word2Vect* mendapatkan hasil accuracy 97.39% pentingnya memperhatikan konfigurasi parameter *multi layer perceptron* untuk mendapatkan hasil terbaik dari parameter yang digunakan. Penelitian yang telah dilakukan menghasilkan beberapa saran untuk penelitian lanjutan, Menambahkan *dataset* yang akan menambah keberagaman dan kompleksitas dalam analisis. Melakukan penanganan pada data yang mengandung kata-kata yang tidak tepat. Penggunaan *Embedding* yang Lebih Kompleks. Mengexplore konfigurasi pada MLP sebagai metode klasifikasi.

Daftar Pustaka

- [1] R. Bagué-Masanés, B. Remeseiro, and V. Bolón-Canedo, "Understanding the Restaurant Recommendations with Feature Selection Techniques with a Personalized Recommender." [Online]. Available: <https://ssrn.com/abstract=4601013>
- [2] Q. Gan, B. H. Ferns, Y. Yu, and L. Jin, "A Text Mining and Multidimensional Sentiment Analysis of Online Restaurant Reviews," *Journal of Quality Assurance in Hospitality and Tourism*, vol. 18, no. 4, pp. 465–492, 2017, doi: 10.1080/1528008X.2016.1250243.
- [3] D. Ekawati and M. L. Khodra, "Aspect-based sentiment analysis for Indonesian restaurant reviews," *Proceedings - 2017 International Conference on Advanced Informatics: Concepts, Theory and Applications, ICAICTA 2017*, 2017, doi: 10.1109/ICAICTA.2017.8090963.
- [4] K. Zahoor, N. Z. Bawany, and S. Hamid, "Sentiment analysis and classification of restaurant reviews using machine learning," *Proceedings - 2020 21st International Arab Conference on Information Technology, ACIT 2020*, 2020, doi: 10.1109/ACIT50332.2020.9300098.
- [5] C. S. Perone, R. Silveira, and T. S. Paula, "Evaluation of sentence embeddings in downstream and linguistic probing tasks," 2018, [Online]. Available: <http://arxiv.org/abs/1806.06259>
- [6] N. Claypo and S. Jaiyen, "Opinion mining for Thai restaurant reviews using neural networks and mRMR feature selection," *2014 International Computer Science and Engineering Conference, ICSEC 2014*, pp. 394–397, 2014, doi: 10.1109/ICSEC.2014.6978229.

-
- [7] A. Zikri, A. Zikri, and S. Agustian, "Penerapan Support Vector Machine dan FastText untuk Mendeteksi Hate Speech dan Abusive pada Twitter," vol. 7, no. 1, pp. 436–443, 2023, doi: 10.30865/mib.v7i1.5408.
- [8] "ANALISIS SENTIMEN TERHADAP KEBIJAKAN."
- [9] IEEE Communications Society. Indonesia Chapter., Universitas Telkom., and Institute of Electrical and Electronics Engineers, *Proceedings, the 2020 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology : July 7-8, 2020, Bali, Indonesia*.
- [10] S. Alkadri *et al.*, "Utilizing a multilayer perceptron artificial neural network to assess a virtual reality surgical procedure," *Comput Biol Med*, vol. 136, Sep. 2021, doi: 10.1016/j.combiomed.2021.104770.
- [11] H. Zulfiqar *et al.*, "Deep-STP: a deep learning-based approach to predict snake toxin proteins by using word embeddings," *Front Med (Lausanne)*, vol. 10, 2023, doi: 10.3389/fmed.2023.1291352.